AD-A011 929

# IMAGE ANALYSIS AND MODELING

Paul A. Wintz, et al

Purdue University

Prepared for:

Rome Air Development Center
Defense Advanced Research Projects Agency

May 1975

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

2V

198032



RADC-TR-75-144
Technical Report
May 1975

# IMAGE ANALYSIS AND MODELING

Purdue University

Sponsored By
Defense Advanced Research Projects Agency
ARPA Order No. 2893



Approved for public release;
distribution unlimited.

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U. S. Government.

Rome Air Development Center
Air Force Systems Command
Griffiss Air Force Base, New York 13441



58

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>RADC-TR-75-144 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>IMAGE ANALYSIS AND MODELING | | 5. TYPE OF REPORT & PERIOD COVERED<br>Quarterly Progress Report<br>1 Nov 74 - 31 Jan 75 |
| | | 6. PERFORMING ORG. REPORT NUMBER<br>TR-EE 75-3 |
| 7. AUTHOR(s)<br>Paul A. Wintz<br>T. S. Huang | | 8. CONTRACT OR GRANT NUMBER(s)<br>F30602-75-C-0150 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Purdue University<br>School of Electrical Engineering<br>West Lafayette IN 47907 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>61101E<br>28930001 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Defense Advanced Research Projects Agency<br>1400 Wilson Blvd<br>Arlington VA 22209 | | 12. REPORT DATE<br>May 1975 |
| | | 13. NUMBER OF PAGES<br>[illegible] |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)<br>Rome Air Development Center (IRRO)<br>Griffiss AFB NY 13441 | | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE<br>N/A |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer:
John M. Trossbach, Capt, USAF
AC 315 330-2353

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image Processing
Digital Image Processing
Image Restoration
Image Coding
Pattern Recognition
Syntactic Pattern Recognition
Image Segmentation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report summarizes the results of the research program on Image Analysis and Modeling supported to the Advanced Research Projects Agency under Rome Air Development Center Contract F30602-75-C-0150 for the three month period 1 November 1974 to 31 January 1975. The objective is to achieve a better understanding of image structure and to use this knowledge to develop improved image models for use in image analysis and processing tasks such as information extraction, image enhancement and restoration, coding, etc. The (Cont'd)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ultimate objective of this research is to form a basis for the development of technology relative to military applications of machine extraction of information from aircraft and satellite imagery of the earth's surface. Work reported here includes: image segmentation and decomposition, edge detection, texture analysis, clustering, feature extraction, coding and restoration.

This report has been reviewed by the RADC Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

This report has been reviewed and approved for publication.

APPROVED: John M Trossbach Jr

JOHN M. TROSSBACH, JR., CAPT, USAF
Project Engineer



Do not return this copy. Retain or destroy.

IMAGE ANALYSIS AND MODELING

Paul A. Wintz
T. S. Huang

Contractor: Purdue University
Contract Number: F30602-75-C-0150
Effective Date of Contract: 1 November 1974
Contract Expiration Date: 31 October 1975
Amount of Contract: $427,020.00
Program Code Number: 5D30
Period of Work Covered: 1 Nov 74 - 31 Jan 75

Principal Investigators: Dr. Paul A. Wintz and
Dr. Thomas S. Huang
Phone: 317 749-2540

Project Engineer: Capt John M. Trossbach
Phone: 315 330-2353



This research was supported by the Defense Advanced Research Projects Agency of the Department of Defense and was monitored by Capt John M. Trossbach, RADC (IRRO) Griffiss AFB NY 13441.

TABLE OF CONTENTS

# IMAGE ANALYSIS AND MODELING

## RESEARCH SUMMARY AND OVERVIEW

In this report we summarize our research activities during the three month period November 1, 1974 to January 31, 1975. Our research program started October 1, 1973. During the first year we concentrated our efforts on basic research in image structure analysis and image modeling. During the second year we shall continue our basic research, but in the meantime also look into several applications in the information extraction area. We hope that on the one hand the results of our basic research could find practical applications, and on the other hand problems arise from practical situations would guide the direction of our basic research.

In basic research, we will emphasize image segmentation and decomposition, edge detection, texture analysis, linguistic and syntactic approach, and feature extraction. The applications we shall look into include: information extraction for the purpose of tracking flying objects, matching of similar but not identical images, and detection of airports in ERTS pictures.

As part of our research program we are putting together a computer system which will be a node on the ARPA computer network. The system is working locally and will be on the ARPANET in the near future. The availability of such a system will affect our research in two respects. First, because the system includes a flexible digital display unit, we can very conveniently study man-machine interactive techniques. Second, because of the existence of large computers (such as the IBM 360/91, the CDC 7600, and the ILLIAC 4) on the ARPANET, we can afford to study algorithms which may require a large amount of computation.

## IMAGE SEGMENTATION

P. A. Wintz, P. H. Chen, and P. Miller

As reported earlier [1], we have developed a computer algorithm (BLOB) to segment an image into regions. This algorithm has been proven very successful in some classification and coding applications. We are currently working on extensions of this BLOB algorithm to make it useful for other problems. In particular:

(i) We are developing an algorithm for segmenting multispectral images into multidimensional ( 3 or higher dimensional) blobs.

(ii) The original BLOB algorithm uses the mean and variance of gray levels to group points into regions. We are looking into the use of other statistics which may be more suitable for other applications.

### REFERENCE

[1] P. A. Wintz and J. N. Gupta, "Image Segmentation," in ARPA semiannual Report, October 1, 1973-April 30, 1974.

# IMAGE MODELING BY DECOMPOSITION

T. S. Huang and J. Burnett

## I. STATEMENT OF THE PROBLEM

We are developing improved image models based on decomposition which could lead to break throughs in image processing.

Although there are many examples of success in image processing, a large number of problems remain unsolved or only partially solved. Consider specifically the case of image restoration where one tries to compensate for the degradations suffered by the image due to camera motion, lens aberrations, atmospheric turbulence, etc., and the ever-present detector noise. Past techniques in image restoration [1] succeeded in reducing image noise but generally failed in sharpening up the blurred edges of objects in the image. The main reason for this failure is that in most of these techniques the image is modeled by a stationary second-order stochastic process with an exponential autocorrelation function [2,3]. It can be shown [4] that sample functions of such stochastic processes are necessarily continuous - they cannot contain sudden jumps. It is therefore not suprising that restoration techniques based on this kind of image model yield fuzzy pictures.

We believe that the key to getting a good restoration lies in the development of an improved image model. Specifically, an image should be decomposed into several parts. Then each part is treated separately according to its own characteristics. As a first attempt, we will consider an image as the sum of a pure-jump process (representing the edges in the image) and a piecewise Markov process. The jumps will divide the image into regions. Each of these regions is modeled by a Markov process with its own parameters. For restoration, Wiener or Kalman filters could be used for the Markovian

regions, while new techniques perhaps have to be developed for the pure-jump process [5,6].

## II. SIGNIFICANCE OF THE PROBLEM

The development of improved image models based on decomposition will not only effect significant progress in image restoration but also have a profound influence on all areas of image processing.

For example, in efficient coding if we should succeed in decomposing the image into edges and background, the two parts could be coded differently - perhaps chain coding [7] for the edges and transform coding [8] for the background. In pattern recognition, suitable decomposition of an image could be a crucial first step in feature extraction. And in computer graphics, image decomposition seems a good approach to the generation and display of continuous-tone pictures-thus the "block world" of the artificial intelligence community could be displayed as edges plus shading. Finally, the search for objective (mathematical) measures of image quality may be made easier if different criteria are used for the different components of an image.

All in all, we believe that study on image modeling based on decomposition could lead to important break throughs in image processing.

## III. PLAN OF RESEARCH

The aim of our research is to develop improved image models based on decomposition and to apply these models to image processing. We are proceeding along the following overlapping steps:

1) Develop mathematical theory for image decomposition. We are particularly interested in decomposing an image into a pure-jump process plus a piecewise Markov process. The pieces of the latter may further be classified into "slowly-changing background" and "textures".

2) Develop techniques of estimating model parameters from a given image which

could be blurred and noisy.

3) Develop restoration techniques for the various components of a decomposed image. Special attention will be paid to the particular characteristics of each component.

4) Apply the image model and the restoration techniques to real-life degraded images.

5) Apply the image model to coding, classification, and information extraction.

## REFERENCES


[1] B. R. Frieden, "Image Restoration and Enhancement," in Digital Picture Processing, ed. by T. S. Huang, Springer-Verlag, 1975.

[2] C. Helstrom, "Image Restoration by the Method of Least Square," J. Opt. Soc. of Amer., 1967.

[3] A. Habibi, "Two-dimensional Bayesian Estimate of Images," Proc. IEEE, July 1972.

[4] E. Wong, Stochastic Processes in Information and Dynamic Systems, McGraw-Hill, 1971.

[5] Rubin, "Regular Jump Processes and Their Information Processing," IEEE Trans. on Inf. Thes., September 1974.

[6] Seagall, "A Martigale Approach to Modeling, Estimation, and Detection of Jump Processes," Ph.D. Thesis, Stanford University, August 1973.

[7] H. Freeman, "Boundary Encoding and Processing," in Picture Processing and Psychopictorics, ed. by B. S. Lipkin and A. Rosenfeld, Academic Press, pp. 241-266, 1970.

[8] P. A. Wintz, "Transform Coding," IEEE Proc., July 1972.

SYNTACTIC PATTERN RECOGNITION

W. J. Kelly

## I. INTRODUCTION

During the past several years there has been increasing interest in Syntactic Picture Recognition or Description. Calling upon the theory of Formal Languages, this approach aims at producing picture descriptions which exploit structural relationships in the picture. Both theoretical development and application have been reported in the literature. A short review and references may be found in section II of this summary.

The immediate goal of this research is to identify concepts and techniques to solve military information extraction problems. Initial work has concentrated on the use of such techniques for edge detection and the identification of airports in ERTS pictures. In particular, the tools include, i) stochastic grammars to handle image smearing and noise, and, ii) web grammars to exploit fully the structural relationships in the picture.

## II. SYNTACTIC PATTERN RECOGNITION AND EDGE DETECTION

The application of language theory to pattern recognition goes back to the early sixties, when it was recognized that many applications involve picture "structure" that was not expressible with conventional techniques [4]. The extensive mathematical formalism of linguistic theory appears to be a natural way of developing this structure. Two approaches are possible, depending on the kind of "primitives" chosen to correspond to the "terminals" of language theory. The first approach defines the primitives quite generally while the second uses elements of a grid as primitives with simple adjacency relationships [2].

The second approach allows relatively simple adaptation of conventional (string-generating) language theory, but the representations may be ambiguous and are seldom "transparent". The first approach allows sophisticated relationships to be described, but the theory must be established independently, and the problem of detecting complex relationships in a picture is non-trivial.

Use of these descriptions to identify patterns - "parsing" in linguistics - presents a major difficulty. Formal linguistic parsing, e.g. a FORTRAN compiler, makes little allowance for small "deviations" in a string, and such deviations will cause rejection. Any useful picture will contain some noise and "smearing", and this may cause the parsing to fail even though the structure is "evident" to a human observer. The problem may be avoided to some extent by extensive pre-processing of the data, but this is expensive, inefficient and may destroy significant information. A promising approach is to associate probabilities with the grammar and language, forming so-called "stochastic" grammars and languages which will allow the successful parsing of patterns which deviate "slightly" from the ideal. Considerable work on stochastic picture grammars has been done at Purdue [3,10].

The most promising syntactic pattern formalisn using "general" primitives is the "web grammar" approach [7]. There are several formulations, but the idea is to describe graph structure where the nodes represent "features" of the picture and the edges represent "relations". The "rewriting rules" replace "subgraphs" with more complicated subgraphs. There have been many results reported which extend string-grammar theorems to web grammar, but much remains to be done [],4,6,8].

One application where these techniques might be profitably employed is in the area of "edge detection" [9]. The general problem is to find that which human observers would call an "edge, and often the "structure" in the picture could be usefully employed. The principal difficulties are on the one hand,

"smearing" of the edges and noise, and on the other, detection of boundaries between areas characterized more by "texture" differences than by grey level differences.

## III. TENTATIVE RESEARCH PLAN

One could continue extending "picture grammar" theory along the lines of phrase-structure grammars. However, it is not evident that the extensions will be applicable to the problems of pattern recognition. Other approaches, such as "semantic" and "transformational grammar" pattern recognition, have been proposed, but little solid theory exists so it is difficult to judge their utility. To gain insight into the applicability of the present theory to a real problem, the "edge detection" area will be investigated. These problems exhibit both structure and uncertainty and thus seem amenable to the tools of web grammars and stochastic grammars. Two aspects may be distinguished. Insofar as the edges are part of a larger "structured" pattern, relationships more subtle than adjancency may prevail and the formalism of web grammars seems appropriate. However, the phenomena of noise and "smearing" are often local and stochastic "global structure" grammars would not seem useful in handling the problem. Moreover, the physical location of the edges in the picture may be significant and then the adjacency relation would be important. One should also consider the processing problem. For maximum efficiency, it is probable that parallel processing will be employed in application involving pattern structure; parts of a pattern may be intelligible only in light of other parts. We propose to exploit this parallel processing capability by employing a "two-level" scheme to detect edges in noise where there is a priori knowledge of ome structure in the pattern. The basic idea is to separate the problem of tecting the edges in noise from that of detecting the edge structure, but improve the efficiency of each solution by transferring information from

one to another. To locate noisy or "smeared" edges, a parser for a stochastic string generating grammar will be used, but its operation will be controlled by a web grammar parser that recognizes the edge structure. This assumes a parallel processing capability so that the two levels may proceed simultaneously and guide each other. The use of stochastic web grammars would not be intended to handle "local" noise and smearing, but rather any such phenomena which have a structural character (e.g. edges obscured by shadows).

Presently Purdue does not have a true parallel processing capability, but the long-range usefulness of the method makes it attractive to employ either a parallel-processing simulator such as PAX [11] or the ILLIAC IV through the facilities of the ARPANET.

These concepts will be developed using data from highly-structured pictures, in o  to avoid irrelevant detail. Pictures of straight-edged 3-D blocks have been employed for previous edge-detection research, thus giving a "benchmark" and the initial syntactic research will concentrate on these.

## IV. IMMEDIATE GOALS

i) Development of a web grammar for a class of "block" pictures. High priority should be given to ease of parsing: i.e. minimizing "backtracking".

ii) Development of a stochastic string grammar to represent smeared or noisy edges of block pictures. This may be an extension of Swain and Fu's work [10].

iii) Consideration of the type and method of information transfer between the two "levels". Presumably the web grammar would modify the probabilities of the stochastic grammar, and each parallel execution of the string-grammar passe would inform the web grammar passes of already "identified" edges.

iv) Consideration of an "initialization" scheme, since the parsers must have something presented to them. This is related to the choice of an edge segment identification scheme. Presumably standard statistical techniques will be employed. The "upper-level" parse would not merely guide the "lower-level" parse but may also search for edges not detected by initialization.

## REFERENCES


[1] J. M. Brayer and K. S. Fu, "Some Properties of Web Grammars," TR-EE-74-19, Purdue University, 1974.

[2] K. S. Fu and P. H. Swain, "On Syntactic Pattern Recognition, Software Engineering, Vol. 2, (ed. Tou), Academic Press, New York, 1971.

[3] T. Huang and K. S. Fu, "Stochastic Syntactic Analysis and Syntactic Pattern Recognition," TR-EE-72-5, Purdue University, 1972.

[4] U. G. Montanari, Separable graphs, planar graphs, and web grammars, Inf. and Cont. Vol. 16, pp. 243-67, 1970.

[5] R. Marisimhan, A linguistic approach to pattern recognition, Dig. Comp. Lab Report #121, Univ. of Illinois, Urbana, 1962.

[6] J. L. Pfaltz, Web grammar and picture description, Comp. Graph, and Image Proc. Vol. 1, pp. 193-220, 1972.

[7] J. L. Pfaltz and A. Rosenfeld, Web Grammars, Pros. 1st Int. Joint, Conf. Art. Instell, Wash. D.C., pp. 609-19, May 1969.

[8] A. Rosenfeld and J. P. Strong, A grammar for maps, Software Eng., Vol. 2, pp. 227-39, (op.cit.)

[9] A. Rosenfeld and M. Thruston, Edge and curve detection for visual scene analysis, IEEE Transactions on Computers, V. C-20, pp. 252-269, 1971.

[10] P. H. Swain and K. S. Fu, Stochastic programmed grammars for syntactic pattern recognition, Pattern Recognition, Vol. 4, pp. 83-100, 1972.

[11] B. Lipkin and A. Rosenfeld, <u>Picture Processing and Psychopictorics</u>, Academic Press, 1970.

# PICTURE STATISTICS AND TEXTURES

C. Myers and O. R. Mitchell

One might consider a picture to be divided into three components: the "low frequency," the edges, and the textures. Much work has been done on the slowly changing properties of pictures and on edge detection, but a majority of the work on texture analysis and classification has been predominately problem-oriented. We are investigating the various properties of each of the many problem-oriented classification methods to determine the best texture classification method for general images and also to determine a useful model of what texture is.

Many different approaches have been taken to texture classification. These techniques and some of their advantages and disadvantages are noted below.

1. Fourier spectra: An image may be represented by its spectral frequency content. It has been shown that different land classifications such as "rural, woods, and urban" have different spatial spectral characteristics which could be used to classify them [1]. The problems with this approach are:
   a. the frequency spectra require a large picture area to be accurate and only the gross properties are measured.
   b. if the textural pattern is slightly irregular, the pattern will get lost in the noise.
   c. the positional or phase information of the picture is lost.
2. Autocorrelation function: This gives a measure of the picture's correlation with itself (i.e., how closely two points are related

to each other). This function has much the same problem as the Fourier spectrum for image classification.

3. Gray level measures: The mean, variance, entropy and other such measures may be calculated for textural areas and the area classified by these measures. Many classes of texture measures for an assortment of problems have been tabulated.[4] Gray level measures have been successful for textural classification.[2,3] The majority techniques still presuppose that the boundaries of the object in question are known. Also, statistical measures of this type don't retain much of the spatial relationship information of the various pixels.

4. Joint probability density function and related measures: Haralick[5] proposed that a joint density function be produced, relating a pixel to another pixel in a given direction and distance away. From the density function he calculates 14 features for each direction and distance chosen which he then uses for classification of the area. This techniuqe has the advantage that the features are spatially dependent (as opposed to those derived from the frequency spectrum) and that a small vector of numbers is used to classify the area. The primary disadvantage is that a large area is needed to calculate the measures and the locality of texture becomes fuzzy. Also, two textures may occur in the area of interest and any classification attempt would prove futile.

5. Local operators and templates: One way to classify images is to apply some local operator to a picture and classify according to the result of the operator. Examples of local operators are the gradient and Laplacian: one applies the operator, thresholds the

output and hopefully has the edges. Rosenfeld used this technique to find the areas between objects of similar mean [7]. His operator was the sum of the absolute value of the difference between a point and its neighbors. This function was then thresholded (anything below a given threshold value was considered a non-edge) and, on a limited class of pictures, he obtained the objects.

The advantage of using local operators is that the textural information will remain localized so that any further classification will be on a local basis. The disadvantage of this method is that one has to choose his operator, and the choices are infinite. The choice of the operator is quite important, and due to the vague characteristics of texture, the choice is not intuitively obvious. Also, the choice of operators will be extremely resolution and magnification dependent.

Each of the above techniques has its advantages and disadvantages. We are trying to combine techniques in hopes of better and more general results. Some present research includes:

1. Haralick[5] chose fixed distances between points for calculation of the joint probability density function (spatial dependence matrix). We are developing adaptive measures which use the autocorrelation function results to control the distance and region for out density function and local operator calculations. This should provide some immunity to changes in resolutions.
2. We have chosen a sample of sixteen basic textures from Brodatz[6]. We plan to calculate basic gray level measures, Haralick's measures, and others as yet undefined measures and determine which of these measures are most predictable and useful. It is hoped this will

also lead to definitions of texture primatives.

3. Rosenfeld has shown that applying one local operator to a picture and thresholding will be fairly successful for certain classes of pictures. We are attempting to use more than one local operator simultaneously (of varying size and type) and classify areas according to the results of these many operators (similar to the multispectral classification problem). Through these techniques, we hope to gain insight into the nature of texture primatives and to what texture is.

## REFERENCES


[1] H. R. Johnston, H. El-Sum, M. Rudin, E. Sikorsky, S. Feigenbaum, "Rapid Cartographic Processing System Study," AD 454086, October 1964.

[2] Y. P. Chen and K. S. Fu, "Recognition of X-ray Pictures, IEEE SMC-4, No. 2, pp. 145-56, March 1967.

[3] J. W. Bacus and E. E. Gose, "Leukocyte Pattern Recognition," IEEE SMC-2, No. 2, pp. 513-536, September 1972.

[4] J. K. Hawkins, "Textural Features for Pattern Recognition," in Picture Processing and Psychopictorics, Academic Press, New York, 1970.

[5] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural Features for Image Classification," IEEE SMC-3, No. 6, pp. 610-621, Nov. 1973.

[6] O. Brodatz, Textures, New York, Dover, 1966.

[7] E. B. Troy, E. S. Deutsch, and A. Rosenfeld, "Visual Textural Analysis, 3" Tech. Rep. TR-145, University of Maryland Computer Science Center, Feb. 1971.

# IMAGE RESTORATION BY RECURSIVE ESTIMATION

A. C. Kak and D. Panda

## I. The Problem

A degraded image can be represented by the equation

$$\begin{aligned} y(i,j) &= s(i,j) \; X \; g(i,j) + n(i,j) \\ &= \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} s(i-k,j-\ell)\, g(k,\ell) + n(i,j) \qquad (1) \\ &= x(i,j) + n(i,j) \quad , \quad 1 \le i \le m \; , \quad 1 \le j \le n \end{aligned}$$

where $s(i,j)$ = original image gray level

$g(i,j)$ = point spread function of the imaging medium

$n(i,j)$ = independent noise of variance $\sigma_n^2$ and mean zero

$y(i,j)$ = distorted image

The goal here is to derive the linear mean square recursive estimate (LMSE) of the image $s(i,j)$ in terms of the observations $y(k, )$, $1 \le k \le m$ , $1 \le \ell \le n$.

## II. Image Model

The original image is assumed to be stationary and has an exponential auto-correlation function with vertical correlation coefficient $\rho_v$ and horizontal correlation coefficient $\rho_h$. It can be shown that $\hat{s}(i,j)$, LMSE of $s(i,j)$ in terms of the signal $s(k,\ell)$ at all points $(k,\ell) \neq (i,j)$, is a linear combination of the 8 neighbors of $s(i,j)$, the coefficient of the vertical, horizontal, and diagonal neighbors, $\alpha_v$, $\alpha_h$, and $\alpha_d$ respectively, are:

$$\begin{aligned} \alpha_v &= \frac{\rho_v}{1+\rho_v^2} \\ \alpha_h &= \frac{\rho_n}{1+\rho_n^2} \qquad (2) \\ \alpha_d &= -\alpha_v \alpha_h \end{aligned}$$

and, $s(i,j) = \hat{s}(i,j) + z(i,j)$, where $E[z(i,j)z(k,\ell)] = 0$ for $(i,j) \neq (k,\ell)$. This can be expressed in vector form as

$$A_H \vec{s}(i) = A_v \vec{s}(i+1) + A_v \vec{s}(i-1) - \vec{z}(i) \tag{3}$$

where

$$A_v = \begin{bmatrix} -\alpha_v & -\alpha_d & & \\ -\alpha_d & \ddots & \ddots & \\ & \ddots & \ddots & -\alpha_d \\ & & -\alpha_d & -\alpha_j \end{bmatrix} \qquad A_H = \begin{bmatrix} -1 & \alpha_h & & \\ \alpha_h & \ddots & \ddots & \\ & \ddots & \ddots & \alpha_h \\ & & \alpha_h & -1 \end{bmatrix}$$

$$\vec{s}(i) = \begin{bmatrix} s(i,1) \\ \vdots \\ s(i,n) \end{bmatrix}, \qquad \vec{z}(i) = \begin{bmatrix} z(i,1) \\ \vdots \\ z(i,n) \end{bmatrix}$$

Rearranging (3) we have

$$\vec{s}(i) = A_v^{-1} A_H \vec{s}(i-1) - \vec{s}(i-2) + A_v^{-1} \vec{z}(i-1)$$

## III. Solution for A Simple Degradation

In Eq. (1) let us consider a special class of point spread function representing degradation due to camera motion, where

$$g(k,\ell) = \begin{cases} 1, & k=0, \quad \ell_o \leq \ell \leq 0 \\ 0, & k \neq 0 \end{cases}$$

then the observation vector is

$$\vec{y}(i) = G\vec{s}(i) + \vec{N}(i)$$

where $G = \begin{bmatrix} g(o,o) & g(o,1) & \cdots & g(o,\ell_o) & & 0 \\ & \ddots & & & \ddots & \\ & & \ddots & & & g(o,\ell_o) \\ & & & \ddots & & \vdots \\ & & & & \ddots & g(o,1) \\ 0 & & & & & g(o,o) \end{bmatrix}$

We expect the recursive LMSE of $\vec{S}(i)$ in terms of $\vec{Y}(i)$ to be

$$\hat{\vec{S}}(i) = C_1 \hat{\vec{S}}(i-1) + C_2 \hat{\vec{S}}(i-2) + \Gamma \vec{Y}(i-1)$$

and hence the error vector is

$$\vec{S}(i) - \hat{\vec{S}}(i) = A_V^{-1} A_H \vec{S}(i-1) - \vec{S}(i-2) + A_V^{-1} \vec{Z}(i-1)$$

$$-C_1 \hat{\vec{S}}(i-1) - C_2 \hat{\vec{S}}(i-2) - \Gamma \vec{Y}(i-1)$$

It can be shown that $C_1$ and $C_2$ can be determined in terms of $\Gamma(i)$ by using orthogonality principle and making $\vec{S}(i) - \hat{\vec{S}}(i)$ orthogonal to $\vec{Y}(k)$ for all $k \neq (i-1)$ which gives

$$\hat{\vec{S}}(i) = A_V^{-1} (I-\Gamma(i)G)^{-1} \{(I-\Gamma(i)G)A_H - \Gamma(1)G\} \hat{\vec{S}}(i-1) - \hat{\vec{S}}(i-2)$$

$$+ A_V^{-1} (I-\Gamma(i)G)^{-1} \Gamma(i) \vec{Y}(i-1)$$

Finally $\Gamma(i)$ can be obtained by making $\hat{\vec{S}}(i) - \hat{\vec{S}}(i)$ orthogonal to $\vec{Y}(i-1)$.

## IV. Future Work

We are working on the following:

(i) Formulation and solution of the restoration problem based on 3 past neighbors instead of 8 neighbors.

(ii) Application of the method to motion degraded images (on a digital computer).

# IMAGE MODELING BY INDEPENDENT-INCREMENT PROCESSES, II

T. S. Huang and J. Uban

## I. INTRODUCTION

As we mentioned in a previous report [1], one way of viewing DPCM is that it is the best coding method for sources which can be modeled by independent-increment processes. Such a source is completely characterized by $p(\Delta)$, the probability distribution of the increments $\Delta_k$'s. In that report we presented some measured $p(\Delta)$'s and described several easily implementable suboptimum variable-length codes for the $\Delta$'s, including Laemmel code ($L_N$-codes), Golomb codes ($G_N$-codes), and Hasler codes ($H_N$-codes), where N is the basic block size. It is the purpose of the present paper to report some further experimental results on the efficiencies of these codes.

## II. EXPERIMENTAL RESULTS

Four pictures (part of an ERTS frame of the Chicago area-Green channel, a low-altitude aerial photo of several ships in a harbor, a portrait, and a crowd scene) were sampled at three different resolutions: 128 x 128, 256 x 256, and 512 x 512 samples per picture. Each sample is quantized to 7 bits. The $\Delta_k$'s ranged from -127 to 127. A set of messages $\{m_k;\ k = 1, 2,\ldots, 255\}$ were identified with the $\Delta_k$'s in the following way:

$$\Delta = 0 \rightarrow m_1$$

$$\Delta = k \rightarrow m_{2k}$$

$$\Delta = -k \rightarrow m_{2k+1}$$

$$k = 1,2,\ldots,127$$

The messages were then coded with various $L_N$-, $G_N$-, and $H_N$- codes in the manner described in Ref. [1]. The resulting bit rates are listed in Tables 1-4.

The lowest bit rate in each code type is boxed. Also shown are the entropies. The entropies are plotted against linear resolution in Fig. 1.

## III. DISCUSSIONS

The $p_k \stackrel{\Delta}{=} Pr(m_k)$ of the ERTS frame, the portrait, and the crowd are roughly exponential, while that of the aerial phote contains a very large $p_1$ plus a more or less negative-power tail. Therefore, we would expect that L- and G- codes are more efficient than H- codes for the first three images and the reverse is true for the last. This is indeedly largely borne out in the experimental data in Tables 1-4.

We may also make the following more specific observations:

1) In all cases, the best G-code is more efficient than the best L-code.

2) While the optimum value for N in the cases of the L- and the G- codes generally changes from image to image and from resolution to resolution, that in the case of the H- codes is almost always 1. Moreover, the bit rate for the $H_1$ - code is in no cases more than 5% above that of the optimum G-code. Therefore, if a single code is to be used for a variety of images, the $H_1$ - code seems a good choice.

3) When the linear resolution of the image increases, the entropy per $\Delta$ decreases roughly along a straight line (Fig. 1).

### REFERENCE

[1] T. S. Huang and J. Uban, "Image Modeling by Independent-Increment Process," in Image Analysis and Modeling, Semiannual Research Report under an ARPA Contract, School of Electrical Engineering, Purdue University, Oct. 1, 1973-April 30, 1974.

| Code / Bits per Δ / Resolution (samples per picture) | 128 x 128 | 256 x 256 | 512 x 512 |
|---|---|---|---|
| Entropy | 3.76 | 3.34 | 2.78 |
| $L_1$ | 6.76 | 5.44 | 4.16 |
| $L_2$ | 5.62 | 4.77 | $\boxed{3.97}$ |
| $L_3$ | $\boxed{5.12}$ | $\boxed{4.68}$ | 4.29 |
| $L_4$ | 5.37 | 5.20 | 5.07 |
| $G_1$ | 6.76 | 5.44 | 4.16 |
| $G_2$ | 5.18 | $\boxed{4.54}$ | $\boxed{3.92}$ |
| $G_3$ | $\boxed{4.90}$ | 4.59 | 4.31 |
| $G_4$ | 5.30 | 5.18 | 5.08 |
| $H_1$ | $\boxed{4.99}$ | $\boxed{4.58}$ | $\boxed{4.11}$ |
| $H_2$ | 5.37 | 5.00 | 4.62 |
| $H_3$ | 5.72 | 5.48 | 5.24 |

TABLE 1. ERTS Image

| Code \ Bits per Δ \ Resolution (samples per pictures) | 128 x 128 | 256 x 256 | 512 x 512 |
|---|---|---|---|
| Entropy | 1.40 | 1.28 | 1.16 |
| $L_1$ | 6.09 | 5.55 | 4.85 |
| $L_2$ | 5.63 | $\boxed{5.27}$ | $\boxed{4.82}$ |
| $L_3$ | $\boxed{5.59}$ | 5.37 | 5.08 |
| $L_4$ | 5.88 | 5.76 | 5.60 |
| $G_1$ | 6.09 | 5.55 | 4.85 |
| $G_2$ | 5.01 | $\boxed{4.74}$ | $\boxed{4.40}$ |
| $G_3$ | $\boxed{4.97}$ | 4.84 | 4.67 |
| $G_4$ | 5.46 | 5.39 | 5.31 |
| $H_1$ | $\boxed{3.68}$ | $\boxed{3.58}$ | $\boxed{3.48}$ |
| $H_2$ | 4.50 | 4.43 | 4.34 |
| $H_3$ | 5.38 | 5.32 | 5.25 |

TABLE 2. Low-Altitude Aerial Photo

| Code \ Bits per Δ \ Resolution (samples per picture) | 128 x 128 | 256 x 256 | 512 x 512 |
|---|---|---|---|
| Entropy | 4.50 | 3.92 | 3.23 |
| $L_1$ | 10.37 | 7.33 | 5.08 |
| $L_2$ | 7.97 | 5.97 | 4.51 |
| $L_3$ | 6.54 | $\boxed{5.32}$ | $\boxed{4.50}$ |
| $L_4$ | $\boxed{5.98}$ | 5.40 | 5.13 |
| $G_1$ | 10.37 | 7.33 | 5.08 |
| $G_2$ | 6.97 | 5.46 | $\boxed{4.36}$ |
| $G_3$ | 5.77 | $\boxed{5.03}$ | 4.50 |
| $G_4$ | $\boxed{5.70}$ | 5.35 | 5.13 |
| $H_1$ | $\boxed{5.88}$ | $\boxed{5.22}$ | $\boxed{4.52}$ |
| $H_2$ | 6.05 | 5.34 | 4.95 |
| $H_3$ | 6.44 | 5.88 | 5.37 |

TABLE 3. Portrait

| Code \ Bits per Δ \ Resolution (samples per picture) | 128 x 128 | 256 x 256 | 512 x 512 |
|---|---|---|---|
| Entropy | 5.63 | 4.92 | 4.12 |
| $L_1$ | 22.10 | 14.36 | 8.83 |
| $L_2$ | 15.78 | 10.65 | 7.00 |
| $L_3$ | 11.52 | 8.31 | 6.07 |
| $L_4$ | 8.99 | 7.12 | $\boxed{5.87}$ |
| $G_1$ | 22.10 | 14.36 | 8.83 |
| $G_2$ | 12.84 | 8.98 | 6.24 |
| $G_3$ | 8.70 | 6.78 | $\boxed{5.43}$ |
| $G_4$ | 7.15 | $\boxed{6.22}$ | 5.58 |
| $H_1$ | 7.27 | $\boxed{6.29}$ | $\boxed{5.36}$ |
| $H_2$ | $\boxed{7.13}$ | 6.36 | 5.61 |
| $H_3$ | 7.39 | 6.71 | 6.11 |

TABLE 4. Crowd Scene

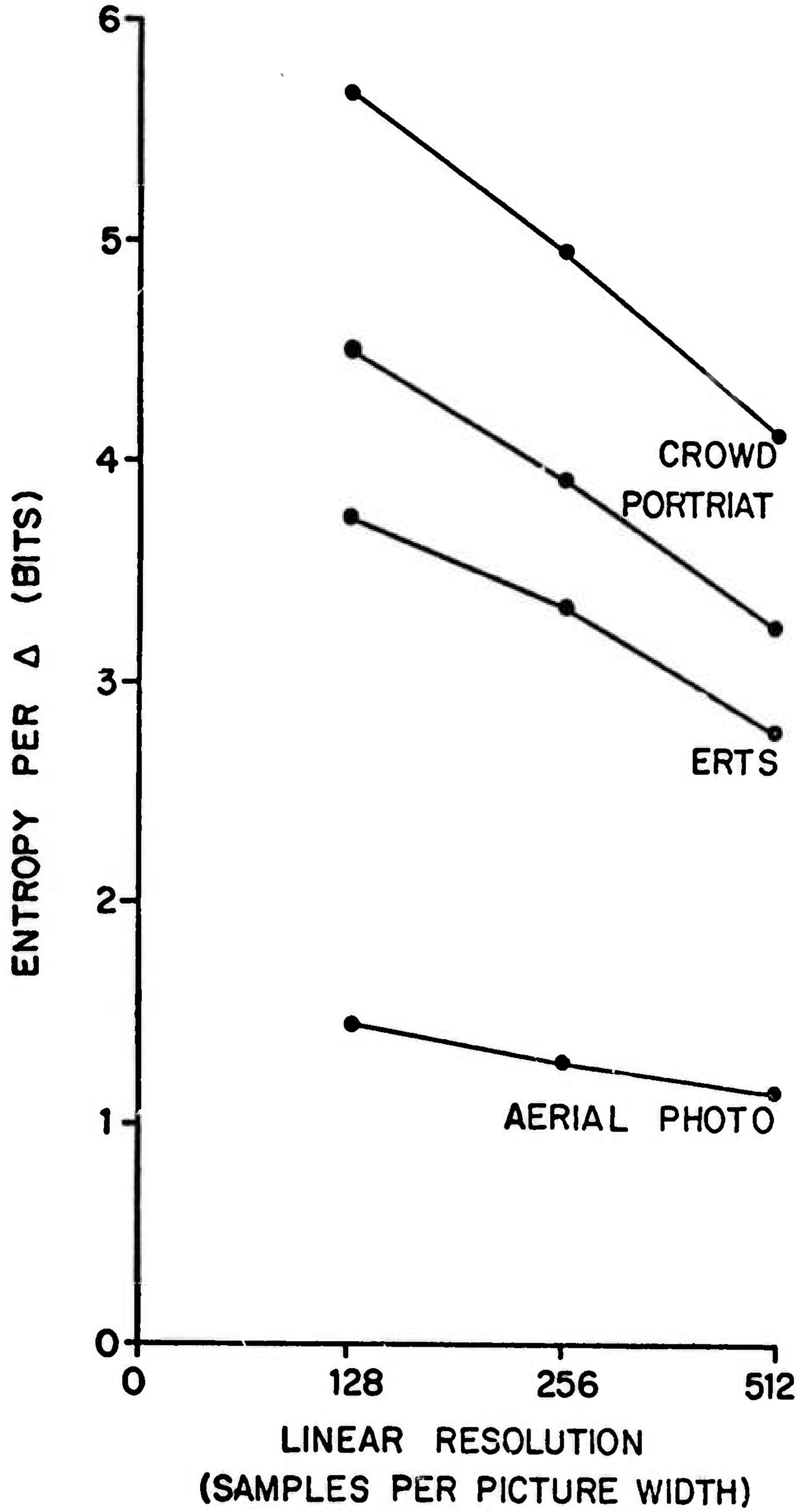


Fig. 1 DPCM Entropy as a Function of Resolution

## SATELLITE IMAGE FORMATION MODEL

O. R. Mitchell and P. L. Chen

### IMAGE FORMATION MODEL

The image of the earth's surface (Chicago area) shown in Fig. 1 has been obtained from the ERTS satellite. As we can see, such images are corrupted by atmospheric effects and cloud formations. These "noise" effects on the signal are not additive in nature but represent multiplication and convolution of noise and signal.

An approach to extracting signals from the multiplied and convolved noise picture has been presented by A. V. Openhein, R. W. Schafer and T. G. Stockham in their paper "Nonlinear filtering of multiplied and convolved signlas." (Proc. IEEE 56, 1264-1291, Aug. 1968).

We are using this approach to remove the effects of clouds and atmosphere in ERTS images. We actually have a three dimensional signal available (2 spatial dimensions plus the four spectral bands representing the third dimension). This opens the possibility of using three dimensional filters to remove the effects of noise from the signal.

The filter design requires that we know the statistics of the noise and the signal. If the noise is produced by cloud cover, the information is most likely available in the spectral signature; if the noise is produced by atmospheric turbulence, these statistics have already been studied and documented. (B. L. McGlamery, Restoration of turbulence-degraded images, J. Opt. Soc. Amer. 57, 293-297, March 1967).

Prior to working with the ERTS images two types of simulated noisy pictures are used as our test samples.

(1) Real physical stimulation:

Images containg multiplicative noise are obtained by taking photographs through either homogeneous or non-homogeneous smoke as shown in Fig. 2.

The power spectrum of the smoke and the power spectrum of the signal are then obtained and an optimum Weiner filter designed for the homomorphic filtering operation to separate the smoke from the signal.

(2) Computer simulation:

A two dimensional pattern of random numbers is generated and filtered to given them the same power spectrum as typical smoke, haze or clouds.

These correlated random numbers are then used to distort digitized images. The dependence of various filtering operations and using different type noises can then be studied.

It is hoped that these studies in filtering and image and noise statistics will enhance our insight into image models.

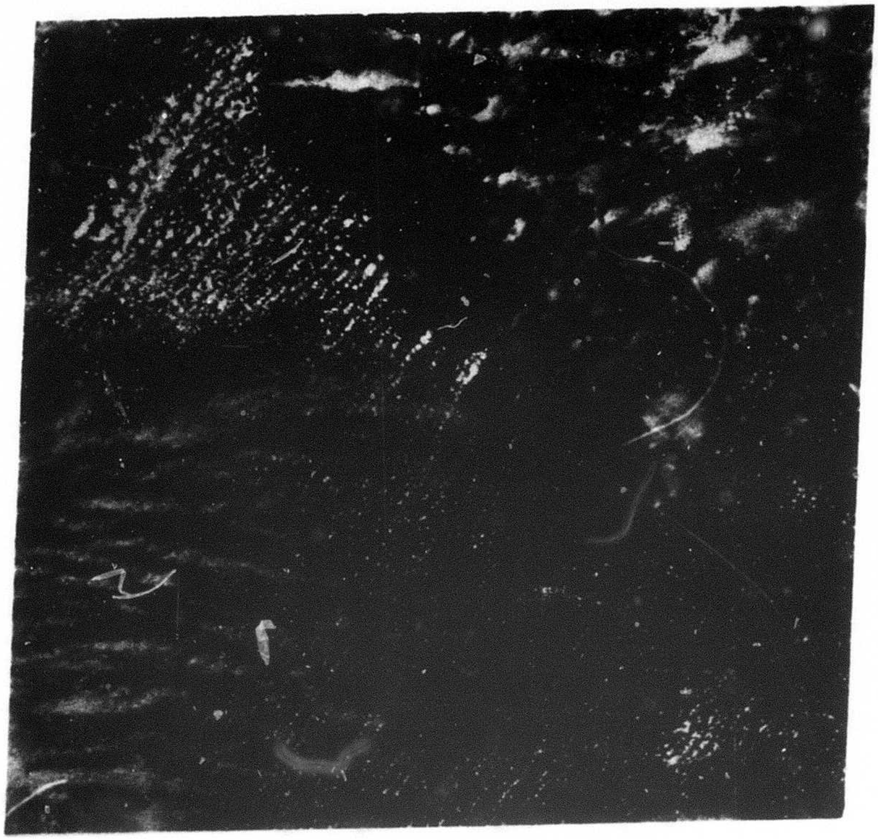

Fig. 1 ERTS Picture with Cloud Cover.

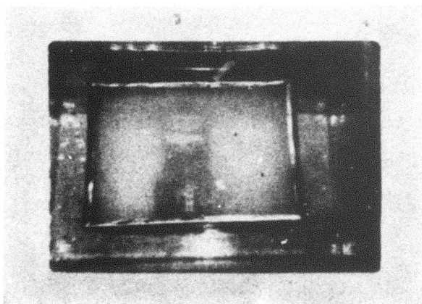

Fig. 2 Laboratory Generated Image with Smoke (multiplicative noise).

ITERATIVE IMAGE RESTORATION, II

T. S. Huang and S. P. Berger

## I. INTRODUCTION

Although much of the work in image restoration has been done in the area of linear spatially-invariant inverse filters, it is evident from past work [1] that in order to restore severly degraded images due to camera motion, lens aberration, atmospheric turbulence, etc., we have to have recourse to linear spatially-varying and nonlinear inverse filters and to make use as much as we can of any a priori knowledge we may have about the original image.

The implementation of linear spatially-varying and nonlinear inverse filters are generally very time consuming. However, the use of iterative methods may provide a relatively efficient solution. In a previous report [2] we have presented an example of restoring linearly degraded image by a particularly attractive iterative technique, the projection method. It is the purpose of the present paper to report some further examples. Experimental results indicated that this method could potentially achieve super resolution in the presence of noise.

## II. EXPERIMENTAL RESULTS

In this section we present two examples of the use of the projection method. In all cases, the only a priori information we made use of was the fact that $f_i \geq 0$.

Our first example is the one-dimensional signal shown in Fig. 1(a). For the purpose of digital processing, 32 samples were taken from it:

$f_0, f_1, \ldots, f_{31}$. We note that the signal consists of two pulses of amplitude 7 separated by a gap of three samples of amplitude 0. This signal was smeared over 10 samples, resulting in the degraded signal $g_i$ (the solid curve in Fig. 1(b)). Specifically,

$$g_i = \frac{1}{10} \sum_{m=0}^{9} f_{i+m}$$

The projection method was applied to this degraded image with an initial guess $\vec{f}^{(0)} = 0$. The results after 1 and 10 cycles of iterations are shown in Fig. 1(b). Note that the pulses were well resolved. Next. noise was added to the smeared signal (the solid curve of Fig. 1(b)) to get a noisy smeared signal (the solid curve of Fig. 2). The noise was independent of the signal, white, and has a uniform distribution with zero mean and a standard deviation $\sigma_n = .5$. The peak-signal-to-rms-noise ratio was then about 23 db. The projection method was applied to this noisy smeared signal with an initial guess $\vec{f}^{(0)} = 0$. The result after 10 cycles of iterations was shown in Fig. 2. Note that the two original pulses were resolved. However, several spurious pulses with smaller amplitudes appeared in the restoration. We note in passing that although the frequency response of the smearing is not strictly bandlimited, the presence of the noise makes it pratically so. Therefore, the result in Fig. 2 indicates that the projection method could potentially achieve super resolution.

Our second example is the image of a cross shown in Fig. 3(a). The digitized image contains 128 x 128 samples. The background is at gray level 255 and the center of the cross at gray level 0. The halftone displays in Figs. 3 and 4 were obtained on a Gould electrostatic printer. This original image was linearly smeared in the horizontal direction. The smearing was spatially-invariant along any particular row. It was, however, different

from row to row. At the center of the image, there was no smearing. The duration of the smearing increased linearly as we moved away from the center. Noise was then added to the smeared image to obtain the noisy smeared images of Fig. 3(b) and 4(a). The noise was independent of the signal, white, and Gaussianly distributed with zero mean. The peak-signal-to-rms-noise ratios for the images in Fig. 3(b) and 4(a) are 36 db and 30 db, respectively. The projection method was applied to these noisy smeared images, with the initial guess equal to the degraded image. The results after 1, 5, and 15 iterations are shown in Figs. 3(c)-(e) and 4(b)-(d). We note that in both cases the image was sharpened up after a sufficient number of cycles of iterations. However, the noise was amplified. Generally, as we do more and more iterations, the image becomes sharper and sharper, but in the meantime the noise becomes more and more amplified. One way to decide on the optimum number of iterations is simply to look at the result after each cycle of iteration, and choose the visually best one. This procedure will be greatly facilitated if we are working on a man-machine interactive system with a fast display unit.

## III. CONCLUDING REMARKS

We have presented some further results of applying the projection method to image restoration. To reiterate, this iterative method has many advantages, specifically:

1. It always converges. When the solution is unique, it will converge to it. When the solution is not unique, it will converge to that solution which is closest to the initial guess solution.
2. It can readily take advantage of a priori information about the image.
3. The computation time requirement is not excessive.

Our experimental results indicated that the projection method could potentially achieve super resolution.

## REFERENCES


[1] B. R. Frieden, "Image Restoration and Enhancement," Chap. 4 of Digital Picture Processing, ed. by T. S. Huang, Springer-Verlag, 1975.

[2] T. S. Huang and D. Barker, "Iterative Image Restoration, in Image Analysis and Modeling," Seimiannual Research Report under ARPA Contract, School of Electrical Engineering, Purdue University, Oct. 1, 1973-April 30, 1974.


## FIGURE CAPTIONS

Fig. 1(a) original signal

(b) smeared and restored signals.

Fig. 2 Noisy smeared signal and its restoration

Fig. 3(a) original image

(b) noisy smeared image (peak-signal-to-rms-noise ratio = 36 db).

(c) restored image (1 cycle of iterations)

(d) restored image (5 cycles of iterations)

(e) restored image (15 cycles of iterations).

Fig. 4(a) noisy smeared image (peak-signal-to-rms-noise ratio = 30 db)

(b) restored image (1 cycle of iterations)

(c) restored image (5 cycles of iterations)

(d) restored image (15 cycles of iterations).

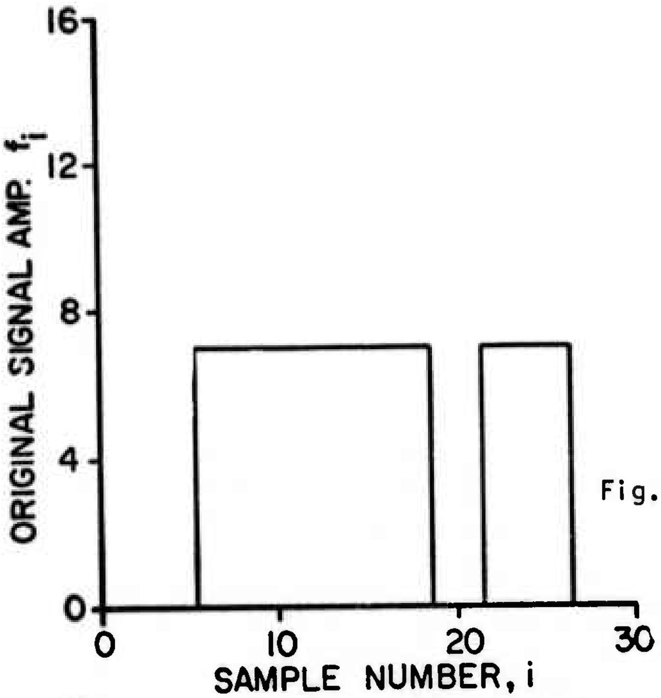


Fig. 1(a) Original Signal.

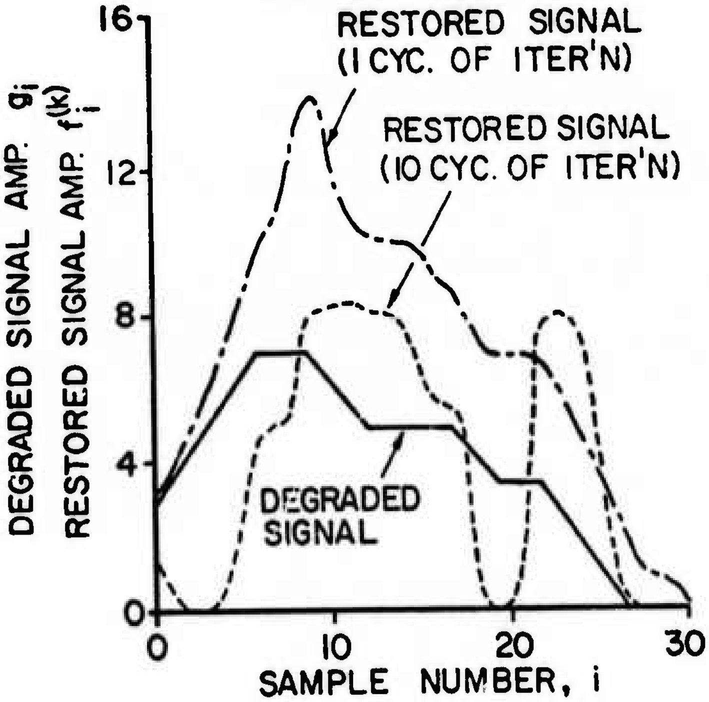


Fig. 1(b) Smeared and Restored Signals.

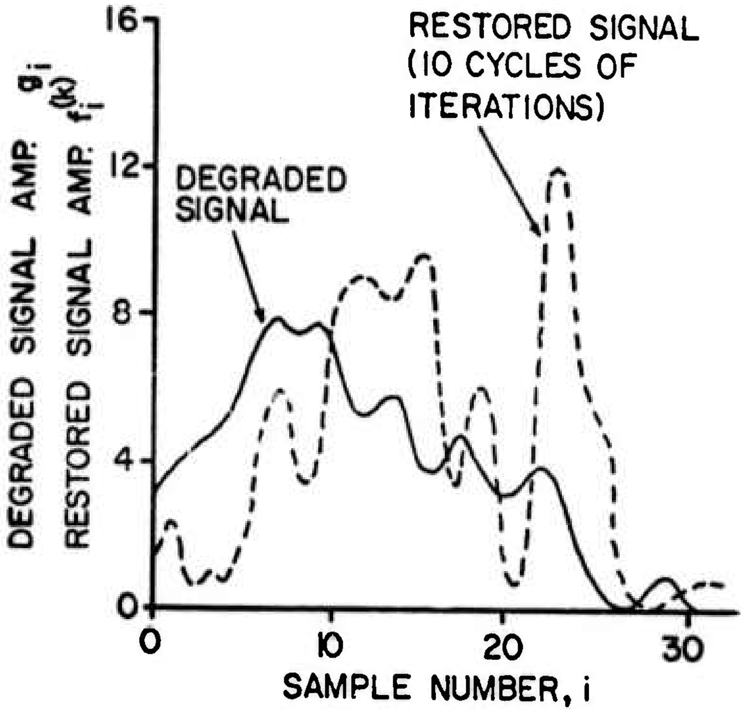


Fig. 2 Noisy Smeared Signal and its Restoration.

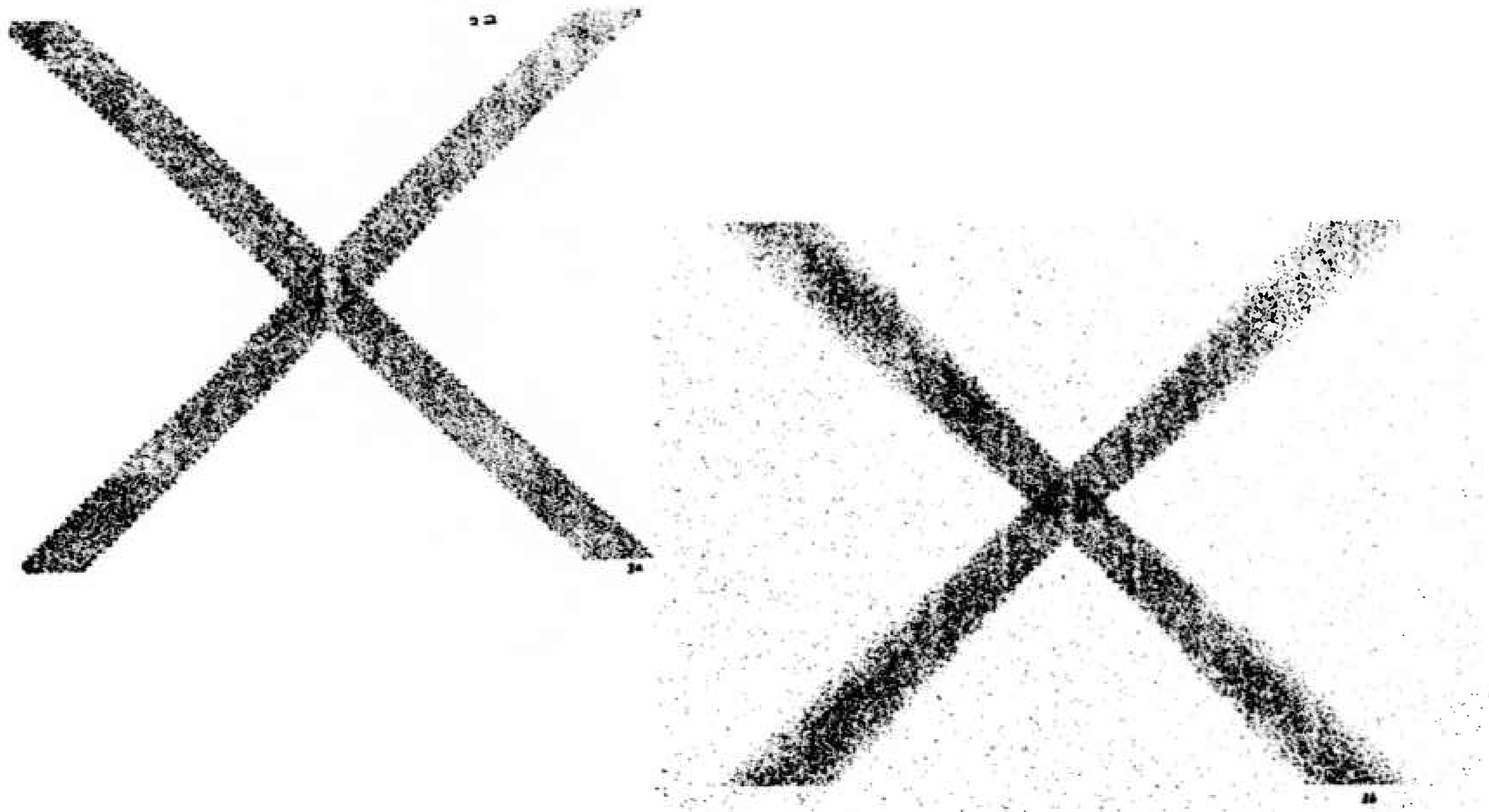

Fig. 3 (a) original image
(b) noisy smeared image (peak-signal-to-rms-noise ratio = 36 db).

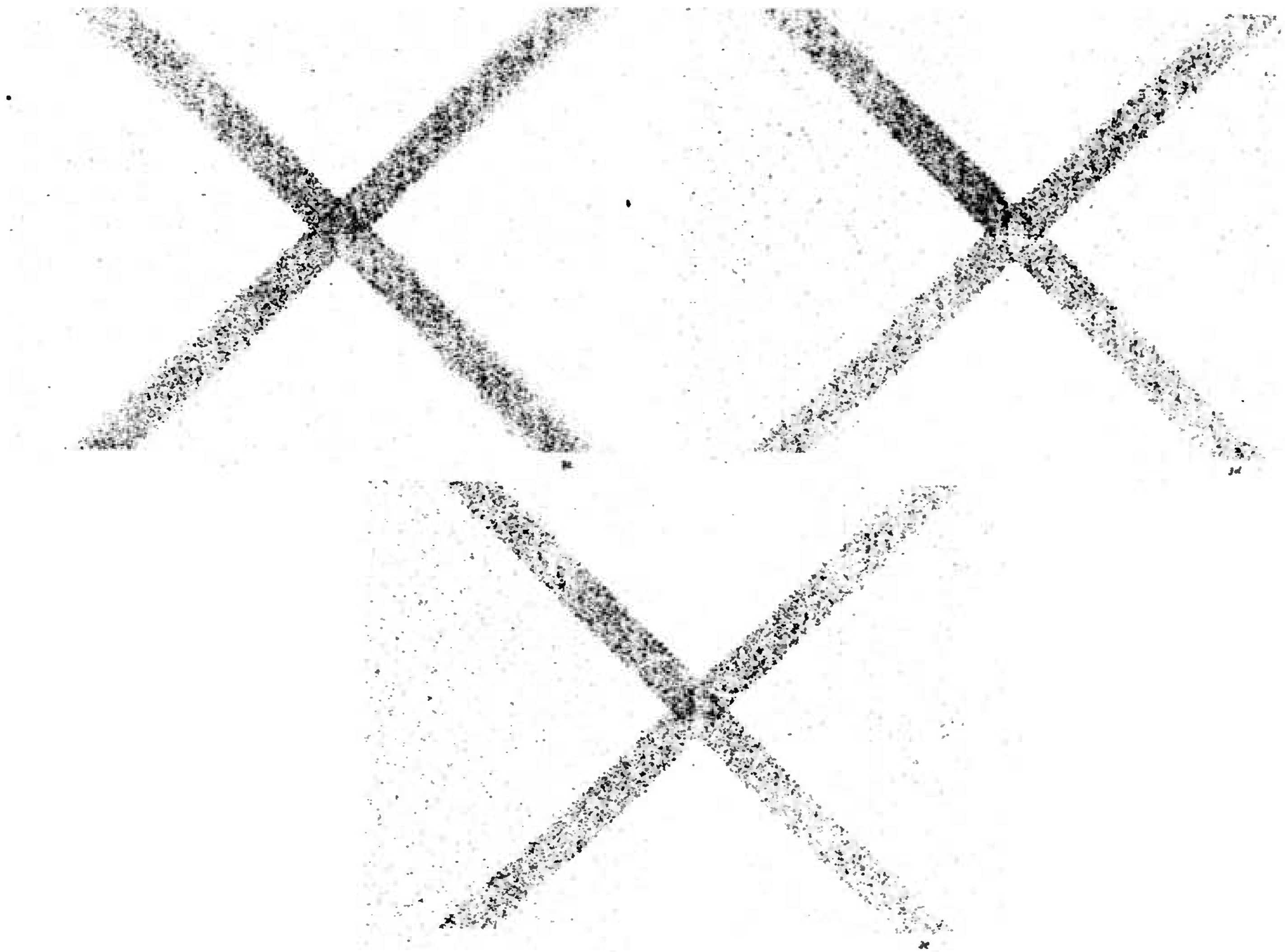

Fig. 3 (cont'd)
(c) restored image (1 cycle of iterations)
(d) restored image (5 cycles of iterations)
(e) restored image (15 cycles of iterations).

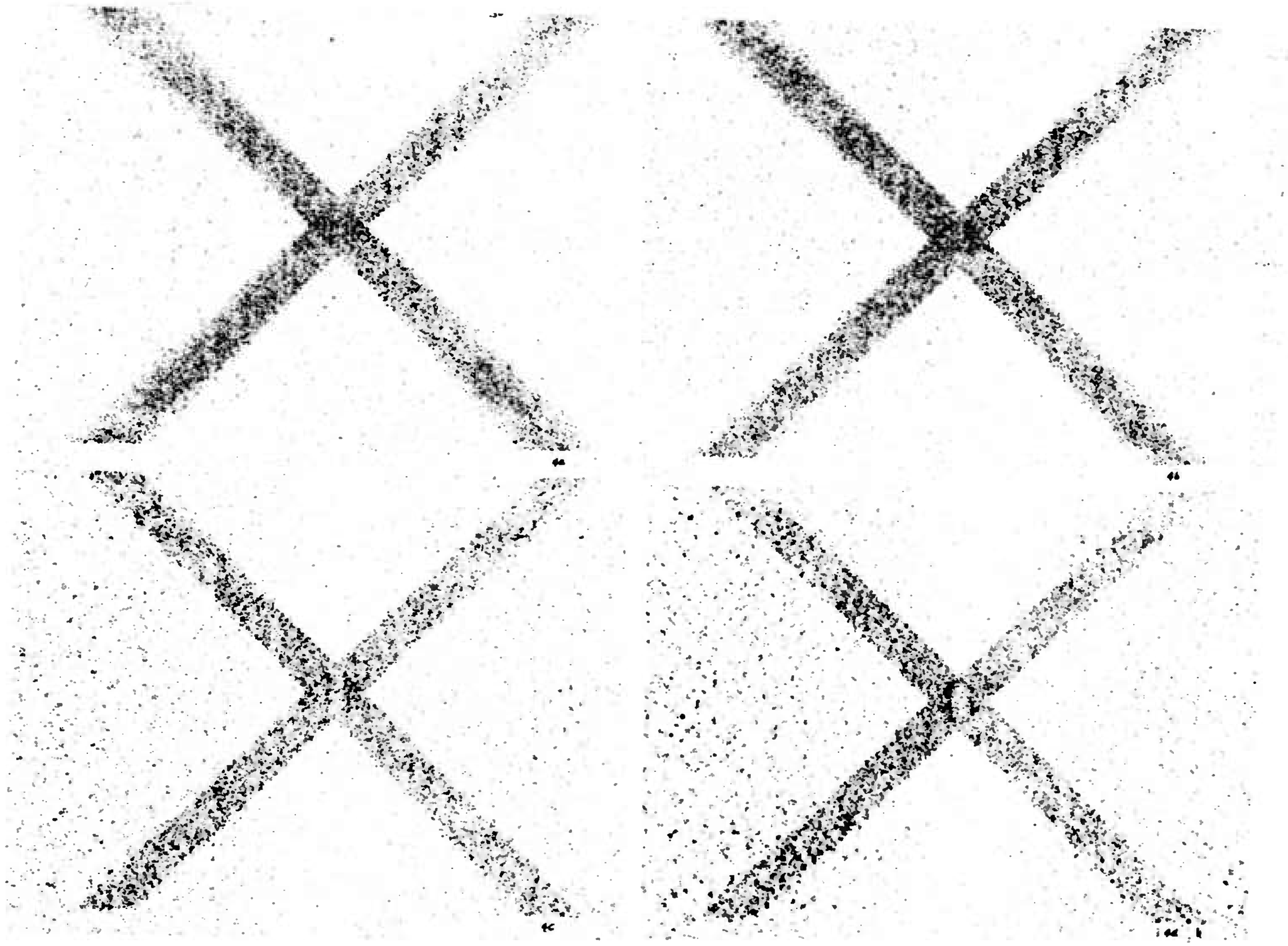

Fig. 4. (a) noisy smeared image (peak-signal-to-rms-noise ratio = 30 db)
(b) restored image (1 cycle of iterations)
(c) restored image (5 cycles of iterations)
(d) restored image (15 cycles of iterations).

# A BRANCH AND BOUND CLUSTERING ALGORITHM*

K. Fukunaga and P. M. Narendra

## I. INTRODUCTION

Clustering may be viewed as a combinatorial optimization problem in that, of the $M^N/M!$ ways in which N objects may be assigned to M classes, the "best" classification, (i.e., one that extremizes the value of a given criterion) is to be determined. One way to find a 'good' classification is to adopt hill climbing techniques which iteratively improve the solution to optimize the criterion. Most clustering algorithms described in the literature [1], [2] belong to this class. While these algorithms are computationally efficient, they do not guarantee a globally optimum solution.

Exhaustive enumeration of all solutions, can, of course determine the globally optimum classification. But the computation required is prohibitive. The method of branch and bound [4] - [7] is a much more reasonable approach to combinatorial optimization problems. In what follows we present a summary of the development of a practical branch and bound clustering algorithm, which finds the globally optimum classification while remaining computationally feasible.

## II. THE BRANCH AND BOUND ALGORITHM

Let $X_1, \ldots, X_N$ be N vectors of dimension n. These vectors are to be clustered into M classes denoted 1, 2, ..., M. Let $\omega_i$ be the class to which $X_i$ is assigned by the clustering, i.e., $\omega_i \in \{1,2,\ldots,M\}$. Let the classification $\Omega$ be defined by

$$\Omega = [\omega_1 \quad \omega_2 \quad \cdots \quad \omega_N]^T$$

*This work was supported in part by the National Science Foundation under Grant GJ-35722.

and $\underline{X} = [X_1^T \ X_2^T \ \dots \ X_N^T]^T$.

The clustering criterion is a scalar function $J(\Omega;\underline{X})$ of $\Omega$ and $\underline{X}$. The clustering problem then is: for a given configuration $\underline{X}$, determine the classification $\Omega^*$, such that

$$J(\Omega^*;\underline{X}) = \underset{\Omega}{\mathrm{Min}} \ J(\Omega;\underline{X}) \tag{1}$$

To seek the above minimum by exhaustive enumeration requires the evaluation of the criterion for $M^N/M!$ classification $\Omega$. However, we may limit the numbers of classification to be enumerated as follows:

Define a partial criterion as

$$J(1,k) \triangleq J(\omega_1,\omega_2,\dots,\omega_k;\ X_1,\dots,X_k) \ , \tag{2}$$

a function of the first k vectors and their classification. Further, suppose that for a given $(\omega_1,\dots,\omega_N)$,

$$J(1,1) \leq J(1,2) \leq \dots \leq J(1,N) \ . \tag{3}$$

Let B be an upper bound on the optimum value of $J(\Omega;\underline{X})$, i.e.,

$$B \geq J(\Omega^*;\underline{X}) \tag{4}$$

Now, if, for a given partial classification: $[\omega_1^a,\ \omega_2^a,\ \dots,\ \omega_k^a]$, $(k \leq N)$, we have,

$$J(\omega_1^a,\ \dots,\ \omega_k^a;\ \underline{X}) > B, \tag{5}$$

then all classifications of the form $[\omega_1^a,\ \omega_2^a,\ \dots,\ \omega_k^a,\ \omega_{k+1},\ \dots,\ \omega_N]$ are suboptimal. This is because,

$$J(\omega_1^a,\ \dots,\ \omega_k^a,\ \omega_{k+1},\ \dots,\ \omega_N) \geq J(\omega_1^a \dots \omega_k^a) > B \ , \tag{6}$$

from (3). Hence, if such an inequality as (5) is satisfied at a certain stage k, $M^{N-k}$ possible trial classifications are found to be suboptimal, and do not have to be enumerated. So the clustering problem reduces to that of finding an efficient way of computing the upper bounds B and enumerating the partial sequences $[\omega_1, \ldots, \omega_k]$, so that the inequality in (5) is satisfied for a large number of the enumerated sequences.

The algorithm is an alternating sequence of partial classification generation and partial classification evaluation steps. All possible classifications are either explicitly evaluated or implicitly rejected. The optimum classification is thereby determined. A detailed account of the algorithm implemented for the within-class scatter criterion may be found in [8].

## III. EXPERIMENTAL RESULTS

The clustering algorithm was applied to 120 samples from a bivariate distribution. Figure 1 shows the optimum classification generated when the algorithm was applied to the data to cluster it into two clusters. Exhaustive enumeration in this case would have required the evaluation of $2^{120}$ different classification. The clustering algorithm obtained the optimal classification in 27 seconds of CPU time on the CDC 6500. This is ample indication that the algorithm is very viable even for large sample sizes.

## IV. CONCLUSION

The theoretical part of this project has been completed. A detailed report will be published [8]. Further work is underway to apply the branch and bound optimization techniques to several problems in Pattern Recognition.

## REFERENCES


[1] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," Proc. of the fifth Berkley Symp., Vol. I, pp. 281-297.

[2] G. H. Ball and D. J. Hall, "ISODATA - A Novel Method of Data Analysis and Pattern Classification," Technical Report, SRI Project 5533, Stanford Research Institute, Menlo Park, Calif., May.

[3] J. J. Fortier and H. Solomon, "Clustering Procedures in Multivariate Analysis," Ed. P. R. Krishnaiah, Academic Press, New York, pp. 493-506, 1966.

[4] S. W. Golomb and L. D. Baumert, "Backtrack Programming," Journ. ACM, Vol. 12, pp. 516-524, Jan. 1965.

[5] A. L. Chernyavskii, "Algorithms for the Solution of Combinatorial Problems based on a Method of Implicit Enumeration," Automation and Remote Control, Vol. 33, pp. 252-260, Feb. 1972.

[6] E. L. Lawler and D. E. Wood, "Branch-and-Bound Methods a Survey," Oper. Res. Vol. 149, No. 4, 1966.

[7] N. J. Nilsson, "Problem Solving Methods in Artificial Intelligence," McGraw-Hill, New York, Chapter 3, 1971.

[8] W. L. G. Koontz, P. M. Narendra, and K. Fukunaga, "A Branch and Bound Clustering Algorithm," to appear in IEEE Trans. on Computers.

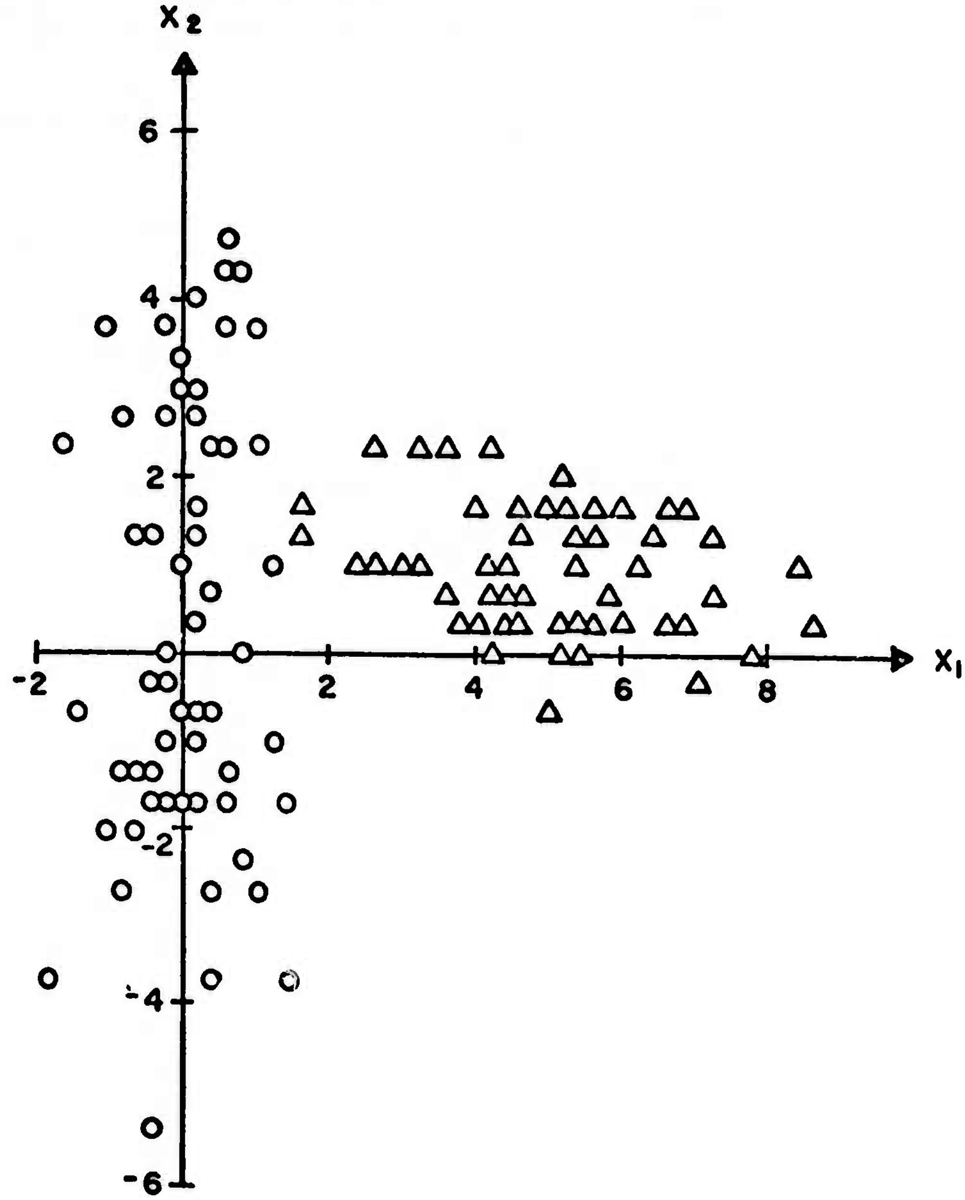


FIGURE 1 SCATTER PLOT OF THE DATA USED IN COMPUTER EXPERIMENTS

APPLICATION OF SINGULAR VALUE DECOMPOSITION*
TO PATTERN RECOGNITION

K. Fukunaga and G. V. Sherman

## I. INTRODUCTION

Classification of sampled images by decision-theoretic methods is usually preceded by a feature extraction technique for reducing the dimensionality of the problem.

Unfortunately most existing statistical methods of feature extraction require the calculation of eigenvectors of the same dimensionality as the original data. Specific examples are factor analysis [1], mixture normalization [2], and the subspace method [3]. In the case of a small 16 x 16 subpicture, the above mentioned eigenvectors would be 256-dimensional; and computation time would be prohibitive.

In this report, we propose to operate on the rows and columns of the picture independently, so that the mathematical manipulations would take place in a vector space of dimensionality equal to that of the rows or columns of the picture.

## II. DESCRIPTION OF THE METHOD

Given the classes $w_k$ $(k=1,\ldots,M)$, we form M standard picture matrices $A^{(k)}$ from the training set $R^{(k)}$ of each class according to

$$A^{(k)} = E_{\underline{A}\in R^{(k)}} \{\underline{A}|w_k\} \tag{1}$$

where E denotes expectation and $\underline{A}$ denotes random pictures. All picture matrices will be assumed to be n x n for simplicity. Next n x n matrices

*This work was supported in part by the National Science Foundation under Grant GJ-35722.

$\Phi^{(k)}$, $\Psi^{(k)}$, and $\Lambda^{(k)}$ (the last being diagonal) are calculated by

$$A^{(k)}A^{(k)^T}\Phi^{(k)} = \Phi^{(k)}(\Lambda^{(k)})^2 \tag{2}$$

$$\Lambda^{(k)} = +[(\Lambda^{(k)})^2]^{1/2} \tag{3}$$

$$\Psi(k) = A^{(k)^T}\Phi^{(k)}(\Lambda^{(k)})^{-1} \tag{4}$$

for $k = 1,\ldots,M$. Let $\underline{A}$ be the random picture matrix which is to be classified. $\underline{K}^{(k)}$ is calculated for $k = 1,\ldots,n$ by

$$\underline{K}^{(k)} = \Phi^{(k)^T}\underline{A}\,\Psi^{(k)} \tag{5}$$

$\underline{A}$ is classified into the class $w_k$ for which classification criterion $J^*(\ell)$ is minimized.

$$J^*(k) = \min_{\ell} \sum_{i=i}^{m} \sum_{j=1}^{m} (\underline{k}_{ij}^{(\ell)} - \lambda_{ij}^{(\ell)})^2 \rightarrow \underline{A} \varepsilon w_k \tag{6}$$

where $\underline{k}_{ij}^{(\ell)}$ is the ijth element of $\underline{K}^{(\ell)}$ and $\lambda_{ij}^{(\ell)}$ is the ijth element of $\Lambda^{(\ell)}$. $m^2$ is the number of features used.

For $m = n$ the classification criterion becomes

$$J^*(k) = \min \operatorname{tr} [(A - A^{(\ell)})\ (\underline{A} - A^{(\ell)})^T] \rightarrow \underline{A}\ \ w_k \tag{7}$$

which is simply the nearest-mean criterion in an $n^2$ dimensional vector space. The success of this method will be demonstrated if a significant reduction in the number of features can be achieved without an excessive increase in classification error.

## III. EXPERIMENTAL RESULTS

The numerals of the Munson multi-author alphanumeric data were used. The second and third sample of each numeral submitted by each author constituted the training set of 98 sampler per class. The first sample of each

numeral submitted by each author made up the test set of 49 samper per class. Of course there are 10 classes corresponding to the 10 numerals. Each numeral contains 24 x 24 points.

In order to investigate the feature extraction capabilities of this technique, the experiment was repeated for 8 values of m. Figure 1 is a plot of classification accuracy versus number of features. Although at first glance 79% classification accuracy seems unimpressive, to the author's knowledge no other group has achieved significantly better results with this data base using such a general approach [4]. Granted ad hoc methods do give somewhat better accuracy, our purpose is to demonstrate a general method applicable to a wide variety of data. Moreover Hussain, et al. have shown that classification errors depend very much on how the training set is selected from this particular data base which further complicates comparisons [4]. Furthermore no effort was made to normalize the size, line width, or rotation of the Munson numerals. Fig. 1 indicates that a reduction from 576 to 144 features, a 75% reduction, causes only a 4% reduction in classification accuracy. Thus the method can indeed be considered successful.

## IV. AREAS FOR FURTHER RESEARCH

Instead of using the SVD, one could expand a picture into the principle components of both rows and columns simultaneously. That is $\Phi$ and $\Psi$ would now be the eigenvector matrices of the autocorrelation matrices of the columns and rows of the $\underline{A}$ respectively. Hopefully a two-dimensional extension of the mixture normalization or subspace method could be developed. Furthermore, the second order statistics of the $\underline{A}$ could be used to predict which of the terms in (6) are small and thus negligible. These approaches are currently being pursued.

## REFERENCES


[1] K. Fukunaga, Introduction to Statistical Pattern Recognition, AC Academic Press, New York, 1972.

[2] K. Fukunaga and W. L. G. Koontz, "Applications of the Karhunen-Loeve Expansion to Feature Selection and Ordering", *IEEE Transactions on Computers*, pp. 311-318, April 1970.

[3] S. Watanabe and N. Pakvasa, "Subspace Method in Pattern Recognition", *Proceedings of the First International Joint Conference on Pattern Recognition*, Washington D. C., pp. 25-32, 1973.

[4] A. B. S. Hussain, G. T. Toussaint, and R. W. Donaldson, "Results Obtained Using a Simple Character Recognition Procedure on Munson's Handprinted Data," *IEEE Transactions on Computers*, pp. 201-205, Feb. 1972.

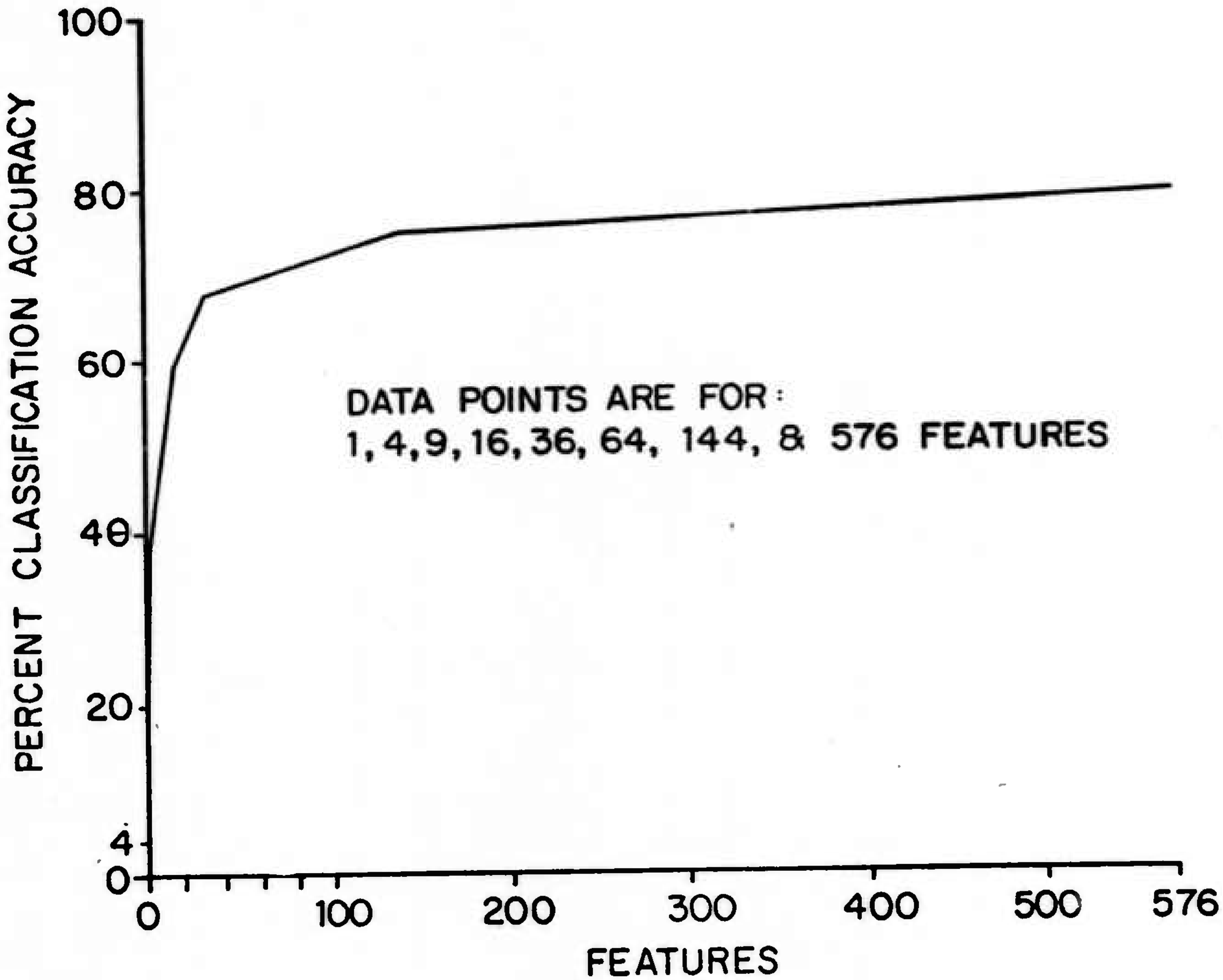


Fig. 1 Classification Accuracy vs. Number of Features.

## FACILITIES

The Purdue IMP arrived and was connected to the net during this research period but interfacing out computer to the IMP was not completed. We are currently debugging the hardware interface between our PDP 11/45 and the IMP. All of our computing over the net was done through a leased 1200 bit/sec telephone line connected to the Illinois ANTS system. Some high volume processing work was done locally due to the low transmission rate of the leased line.

We expect to complete our IMP interface debugging in the immediate future. We are simultaneously developing the use of both ELF and ANTS operating systems on our machines. Present facilities available to us include:

- 1 PDP 11/45 computer; 32K core; disk operating system software
- 1 20-million word disk
- 2 9-track tape drives; 800 and 1600 bpi densities
- 1 7-track tape drive; 556 and 800 bpi densities
- 3 CRT terminals
- 1 card reader
- 1 paper tape reader
- 1 line printer; 600 lines per minute, impact type
- 1 electrostatic graphics hard copy system, 200 points per inch
- 1 video display with disk storage; 3 color refreshed display, 512 x 512 points, 8 bits/color/point; graphics overlay and interactive capabilities
- 1 flying spot scanner

A few hardware interfaces between the PDP 11/45 and the peripherals are yet to be completed. We predict that basic network data transfer capabilities will be available in early - 1975 and complete hardware and software capabilities will be available in mid - 1975.

## SOFTWARE DEVELOPMENT FOR IMAGE PROCESSING

We are developing an integrated collection of picture processing programs with a sensible mix of flexibility and standardization for processing over the ARPA net. The file structure chosen for large images is compatible with that used by the Laboratory for the Application of Remote Sensing (LARS). This allows us to obtain ERTS data directly from LARS and to use the classification programs already developed by LARS.

Features presently incorporated in our software include:

1. Fast Fourier Transform (128 x 128) maximum picture size)
2. Autocorrelation
3. Histogram, joint probability density function
4. Convolution
5. Template matching
6. Utility programs
7. Graphic output
    a. Three-dimensional plots
    b. Contour plots
    c. Grey-level picture on electrostatic printer

Work in progress and future plans include:

8. FFT for 512 x 512 pictures
9. Two-dimensional filtering
10. Picture Editing
11. Histogram Equilization
12. Picture creation and manipulation by use of very simple commands
13. Clustering algorithms
14. BLOB algorithms for image segmentation.

## STAFF

### CO-PRINCIPAL INVESTIGATORS

P. Wintz
T. Huang

### PROFESSORIAL

G. Cooper
K. Fukunaga (on leave)
D. Landgrebe
A. Kak
O. Mitchell

### POSTDOCTORAL RESEARCHERS

M. Kaveh
W. Kelly

### GRADUATE RESEARCHERS

S. Berger
J. Burnett
D. Chan
Wm. Chan
P. H. Chen
P. L. Chen
R. Clark
C. Myers
P. Narendra
B. O'Connor
D. Panda
A. Salahi
G. Sherman
L. Stanfield
G. Tang
E. Wiswell
M. Yoo

### RESEARCH ENGINEER

W. Robey

### UNDERGRADUATE RESEARCHERS

J. Besemer
C. Buckwacter
M. DeMoney
R. Johnson
P. Miller
J. Schwab
J. Uban

### ELECTRONIC TECHNICIANS

D. Azpell

### SECRETARIES

M. Barbour
M. Macha

## JOURNAL PUBLICATIONS


HUANG, T., "Coding of Multilevel Graphics," to appear in IEEE Trans. on Communications, (with H. Koller).

HUANG, T., "Iterative Image Restoration," to appear in Applied Optics, (with D. Baker and S. Berger).

WINTZ, P. A., "On the Effect of Timing Errors in Runlength Codes," to appear in IEEE Trans. on Communications, (with L. C. Wilkins).

WINTZ, P. A., "A Boundary Finding Algorithm and its Applications," to appear in IEEE Trans. on Circuits and Systems, (with J. N. Gupta).

FUKUNAGA, K., "The Estimation of the Gradient of a Density Function and its Applications in Pattern Recognition," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "k-Nearest Neighbor Bayes Risk Estimation," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "A k-Nearest Neighbor Mean-Square-Error Approach to Bayes Risk Estimation," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "Unbiased Estimation of Polynomial Functions Involving Probability Densities," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "Composite Goodness of Fit Tests for Distributions with Complete Sufficient Statistics," to appear in IEEE Trans. on Information Theory, (with D. L. Kessell).

FUKUNAGA, K., "A Test for Multivariate Normality with Unspecified Parameters," to appear in IEEE Trans. on Information Theory, (with D. L. Kessell).

FUKUNAGA, K., "Pseudo Eigenvectors and their Applications to Pattern Recognition," to appear in IEEE Trans. on Electronic Computers, (with G. V. Sherman).

FUKUNAGA, K., "A Branch and Bound Clustering Algorithm," to appear in IEEE Trans. on Computers, (with P. M. Narendra).

FUKUNAGA, K., "A Branch and Bound Algorithm for Computing the k-Nearest Neighbors," to appear in IEEE Trans. on Computers, (with P.M. Narendra).

## CONFERENCES AND WORKSHOPS

WINTZ, P. A., "Image Pattern Recognition for Satellite On-Board Processing of ERTS Resources Data," presented at Pattern Recognition Workshop co-sponsored by NSF and Government Products Div., EIA University of Maryland, Silver Spring, MD., November 1974.

WINTZ, P. A., "Purdue Image Processing," presented at IEEE Conference on Decision and Control, Phoenix, Arizona, November 1974.

WINTZ, P. A., "Earth Resources Data Processing," Seminar presented at the University of Tennessee, Knoxville, Tennessee, November 1974.

HUANG, T., "Easily Implementable Suboptimum Codes," accepted for presentation at the 1975 ICC, San Francisco, CA.

## BOOKS

HUANG, T., (ed), Digital Picture Processing, to be published by Springer-Verlag, 1975.

KAK, A.C., Digital Picture Processing, by A. Rosenfeld and A. C. Kak, Academic Press, Spring 1975.

# MISSION
## of
## Rome Air Development Center

*RADC is the principal AFSC organization charged with planning and executing the USAF exploratory and advanced development programs for information sciences, intelligence, command, control and communications technology, products and services oriented to the needs of the USAF. Primary RADC mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, and electronic reliability, maintainability and compatibility. RADC has mission responsibility as assigned by AFSC for demonstration and acquisition of selected subsystems and systems in the intelligence, mapping, charting, command, control and communications areas.*